

Input a data set  $D$  consisting of text documents, each labeled as belonging to 0 or more classes from a set or hierarchy of classes  $S$ .

110

Construct a single vector representation of text features extracted or associated with each text document in  $D$ .

120

For each labeled text document in data set  $D$ , create a training set  $T(D)$  by labeling the vector by the same set of classes used to label the text documents.

130

For each labeled text document in data set  $D$ , create a to induce classification methods that can be used to assign classes in  $S$  to a hitherto unseen feature vector with the same structure as those in  $T(D)$ .

140

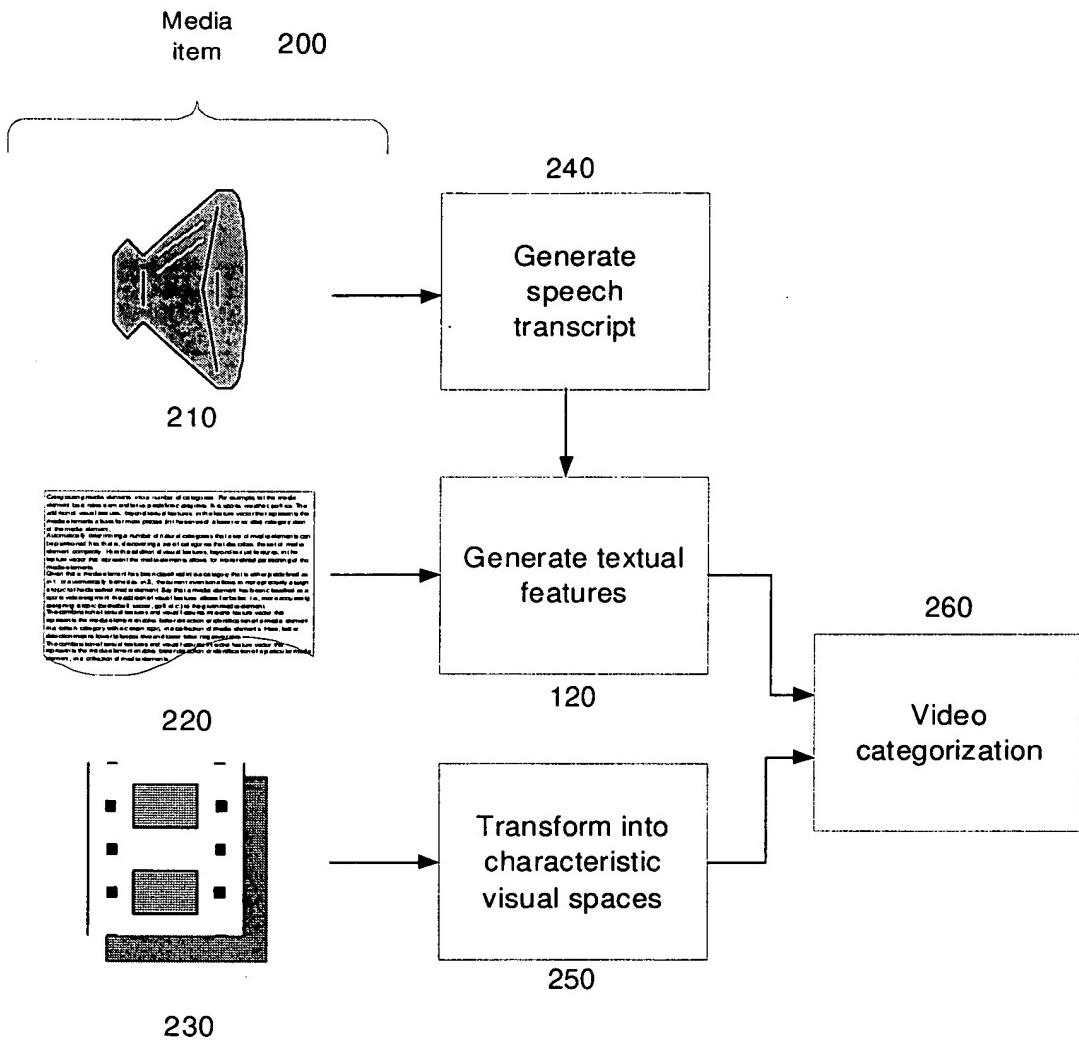
For each class in  $S$ , output the classification methods that can be used to assign that class to a hitherto unseen text document by applying the methods to a feature vector derived from there text document in the same way that the feature vectors in  $T(D)$  were derived from the data set  $D$ .

150

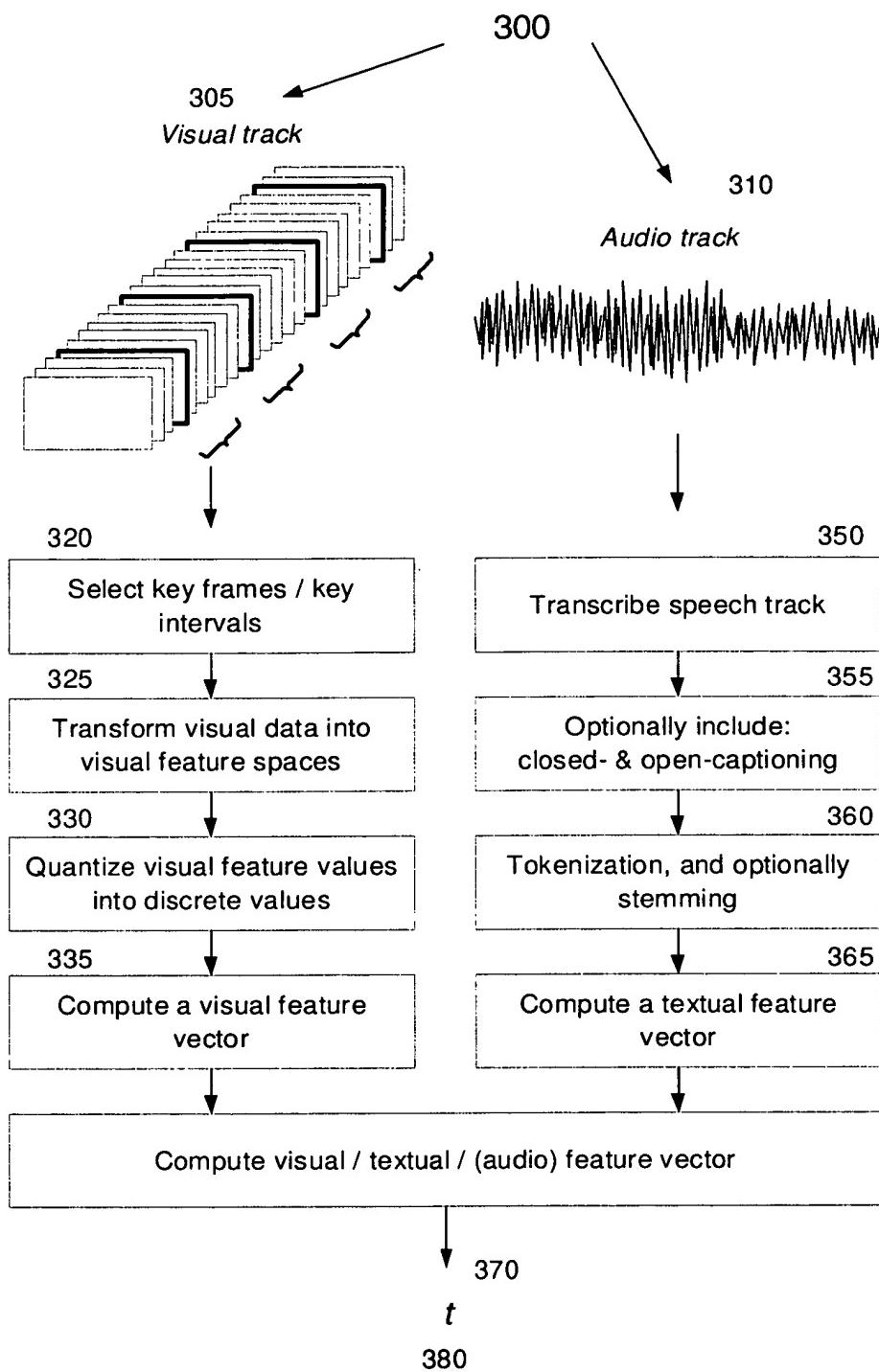
## Prior Art

100

Figure 1



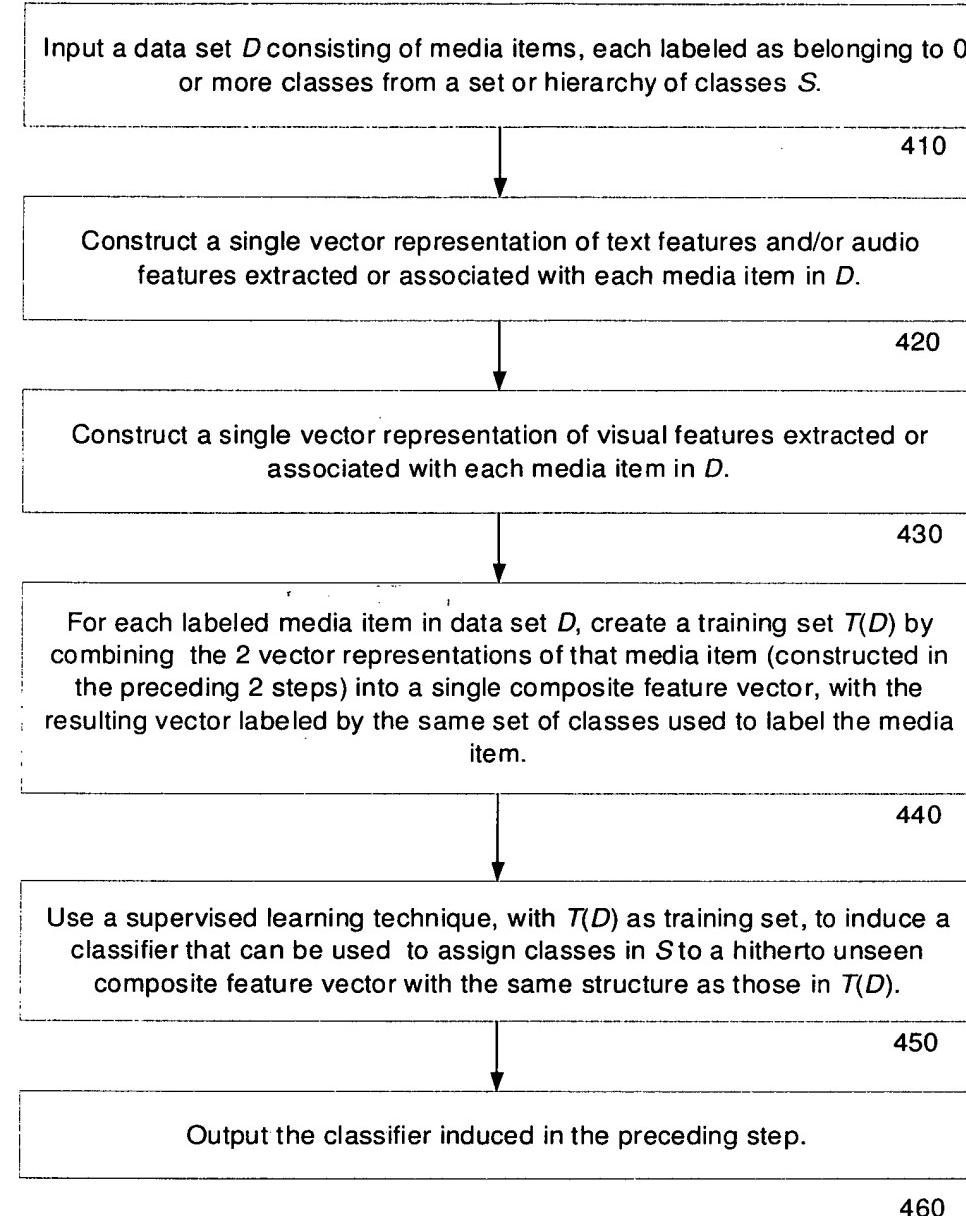
**Figure 2**



**Figure 3**

00000000000000000000000000000000

400



**Figure 4**

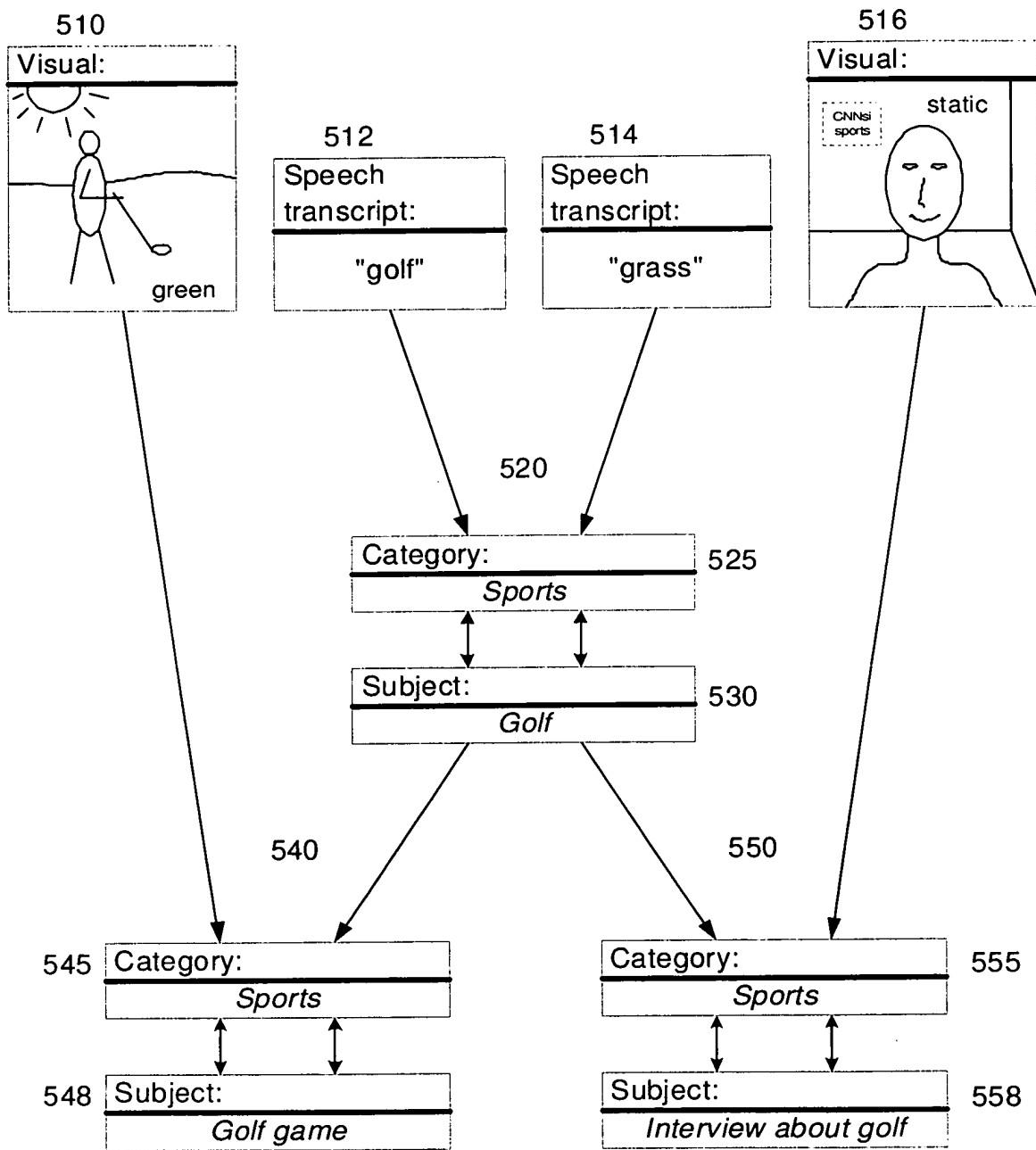
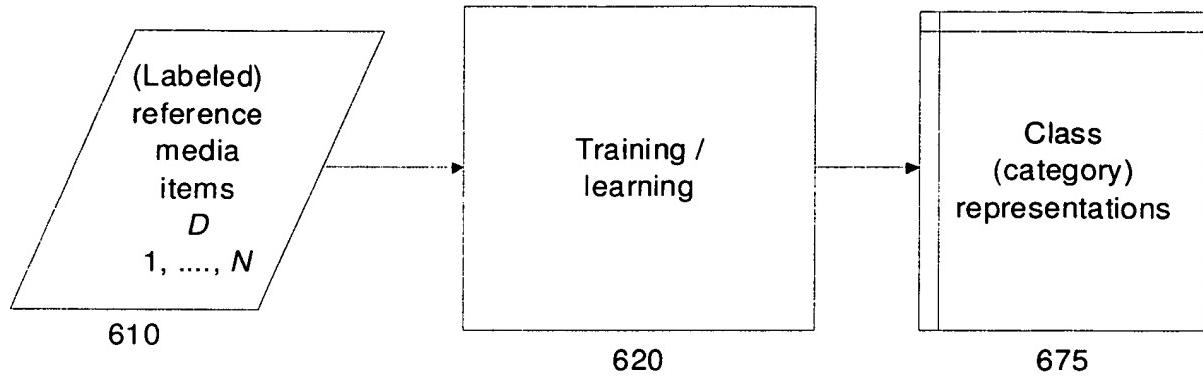
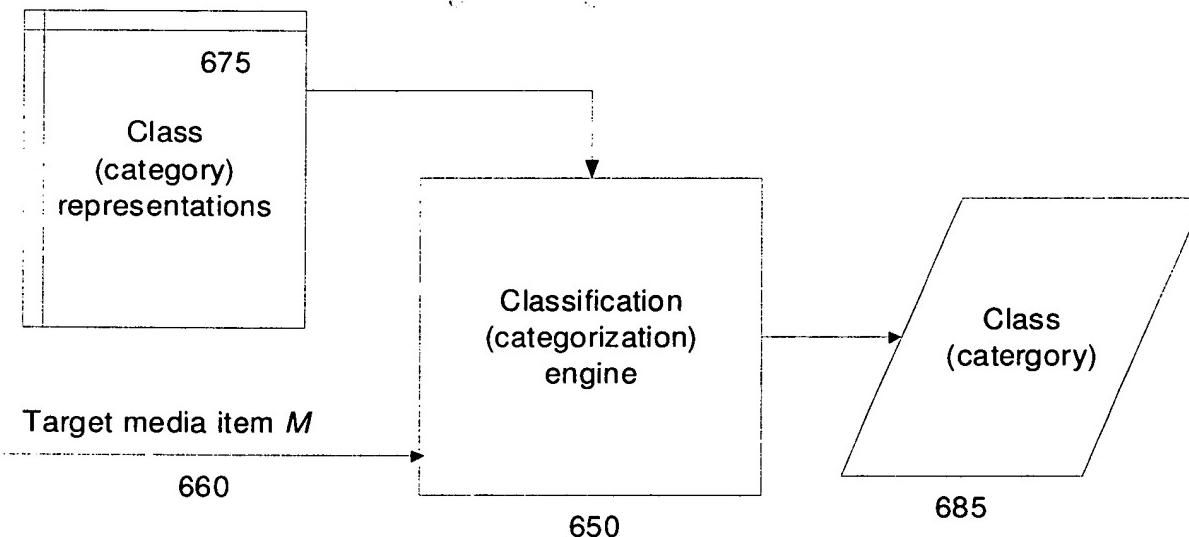


Figure 5



Training / learning phase 610

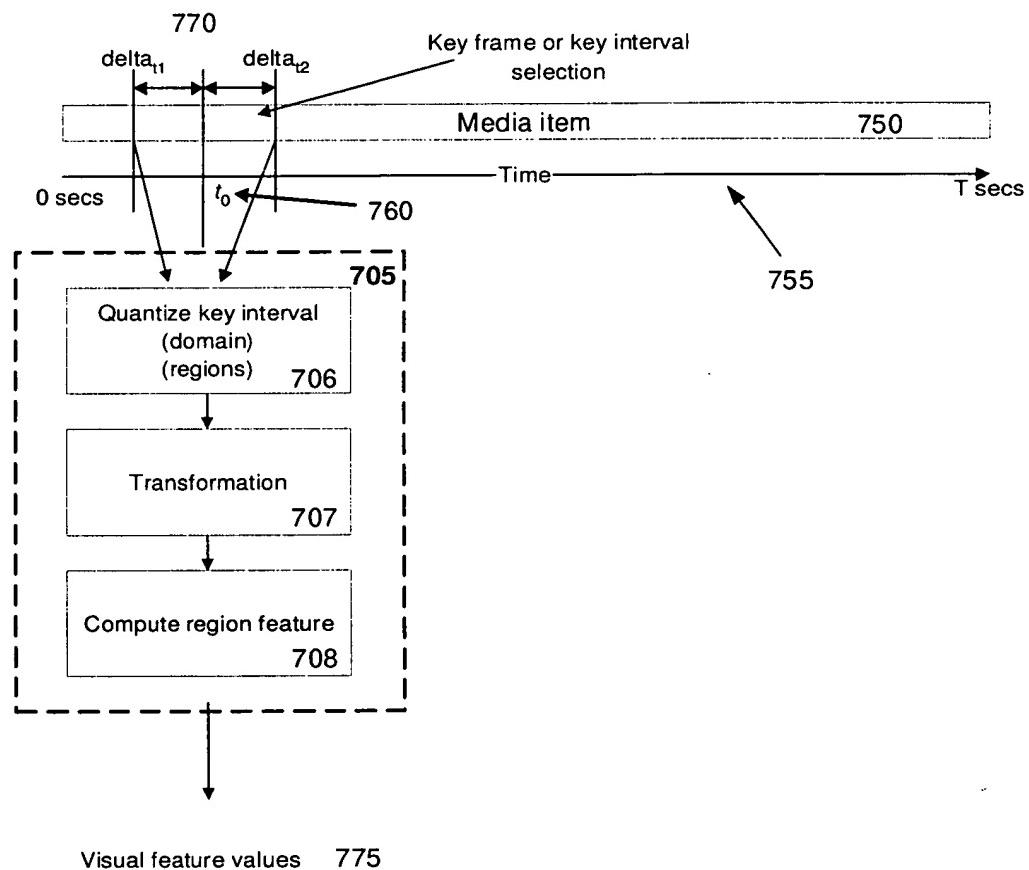
**A**



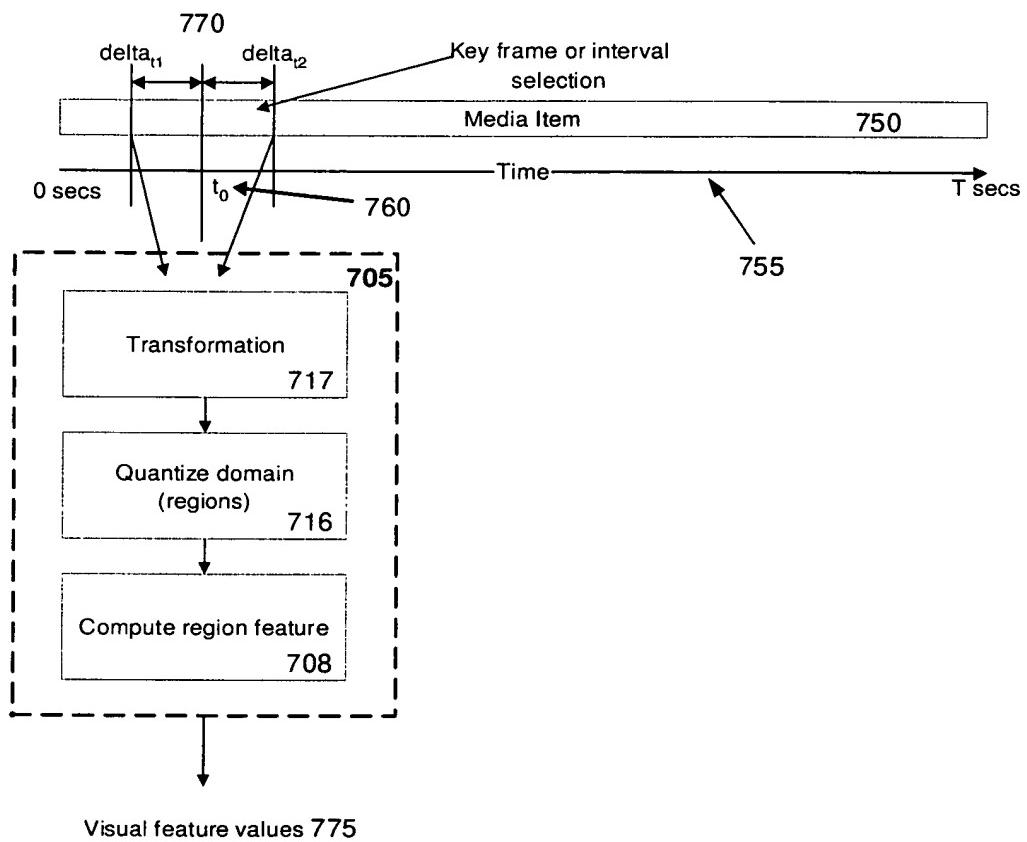
Classification phase 615

**B**

**Figure 6**

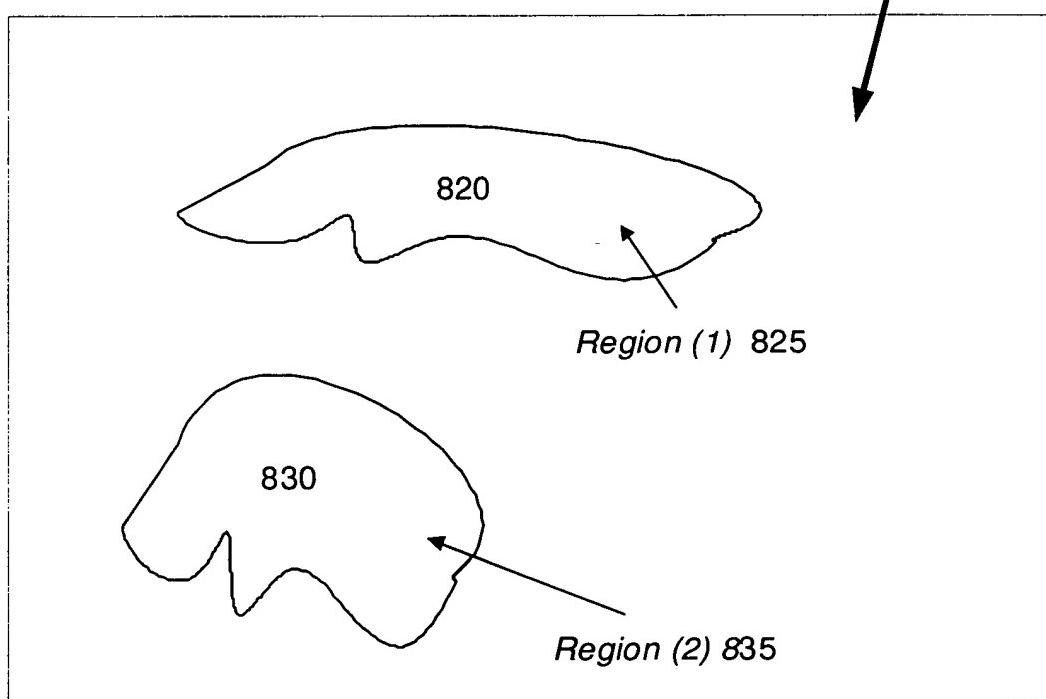


**Figure 7A**

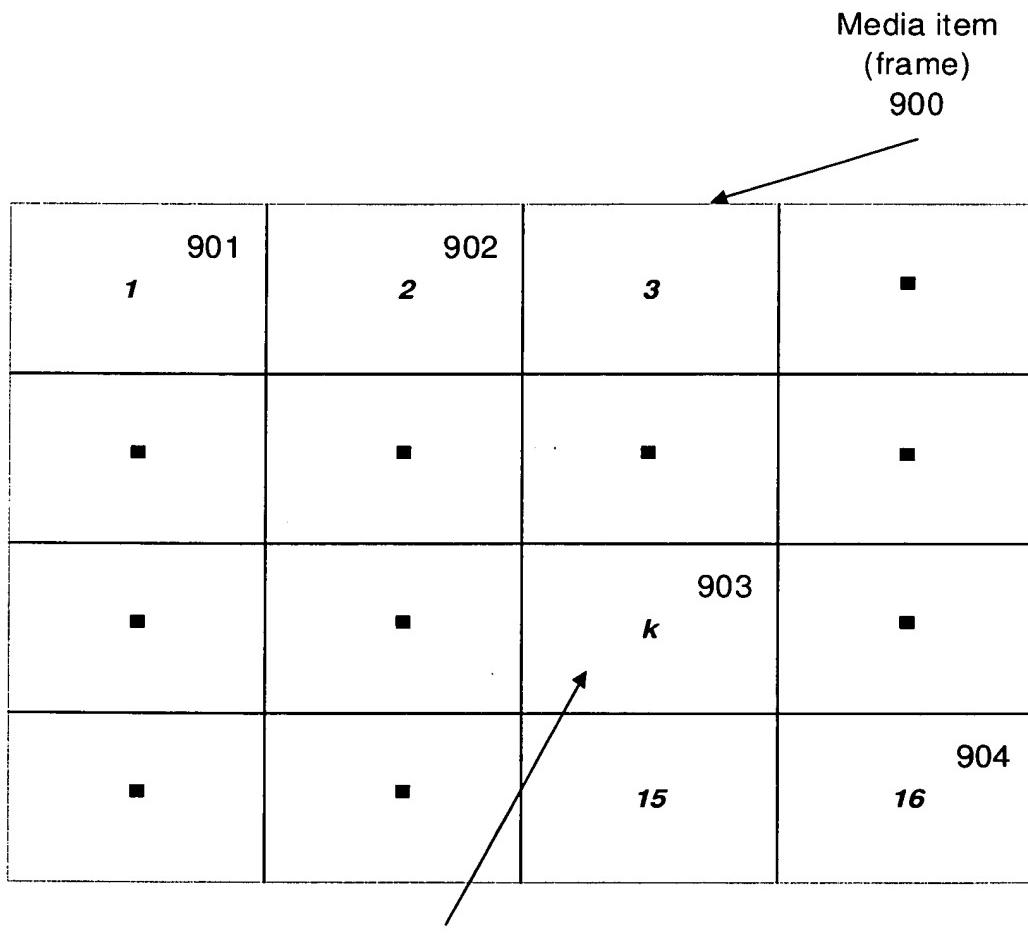


**Figure 7B**

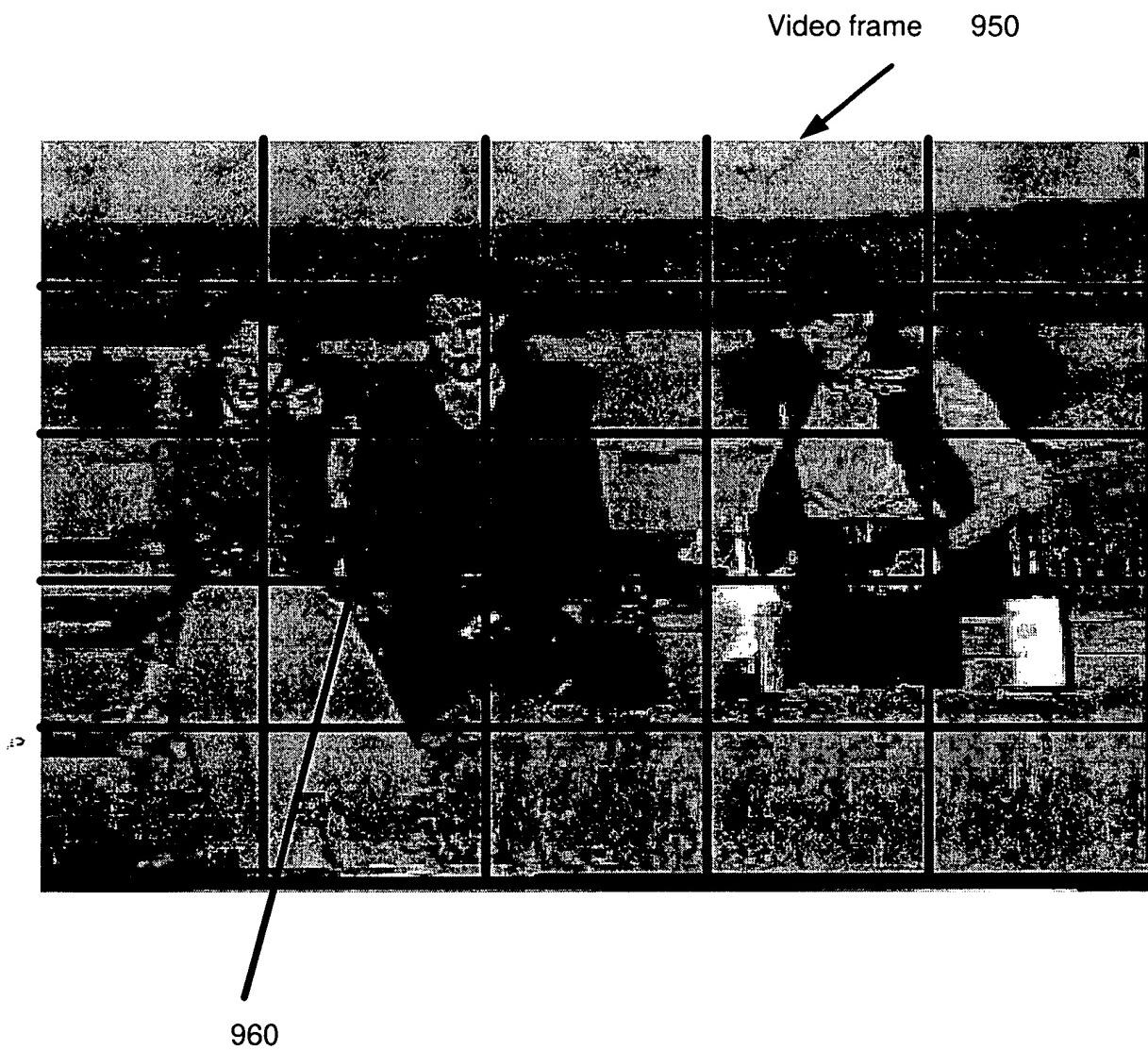
(Part of)  
media item  
810



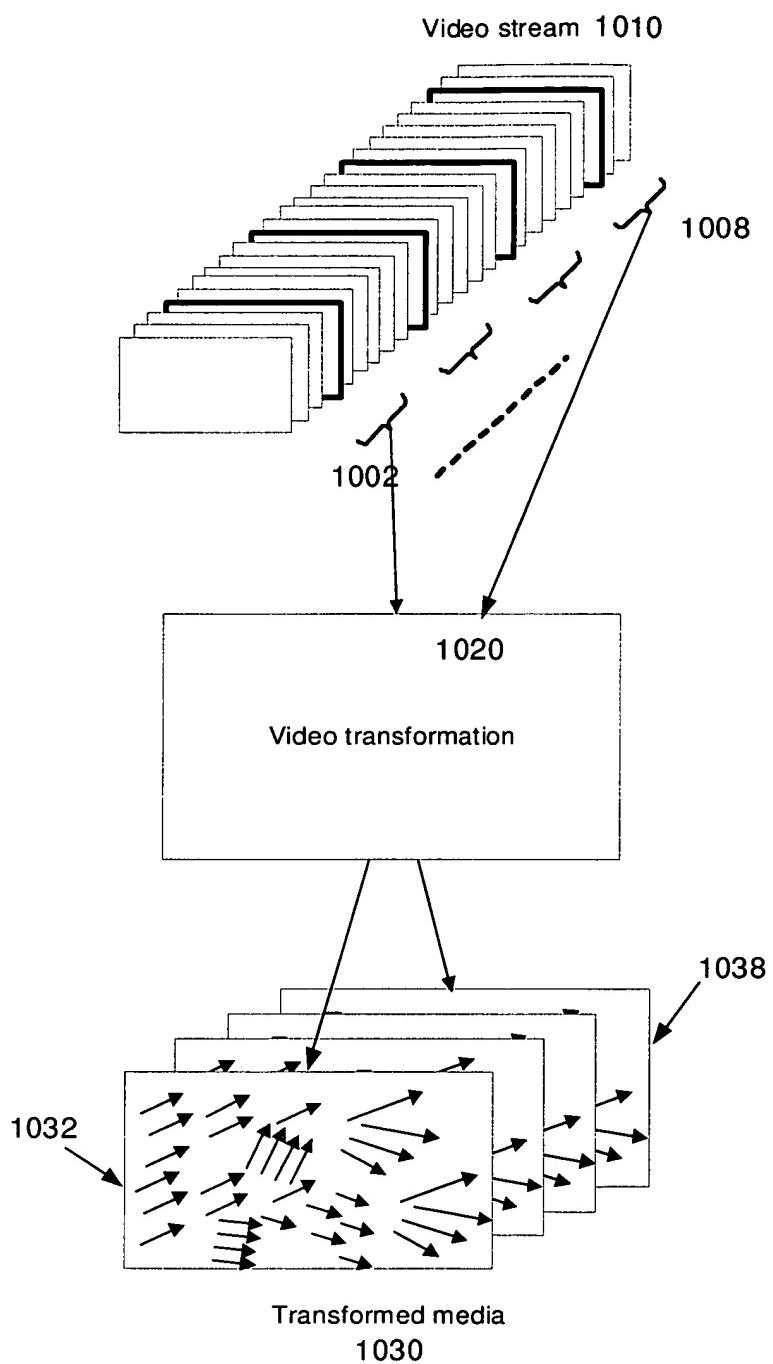
**Figure 8**



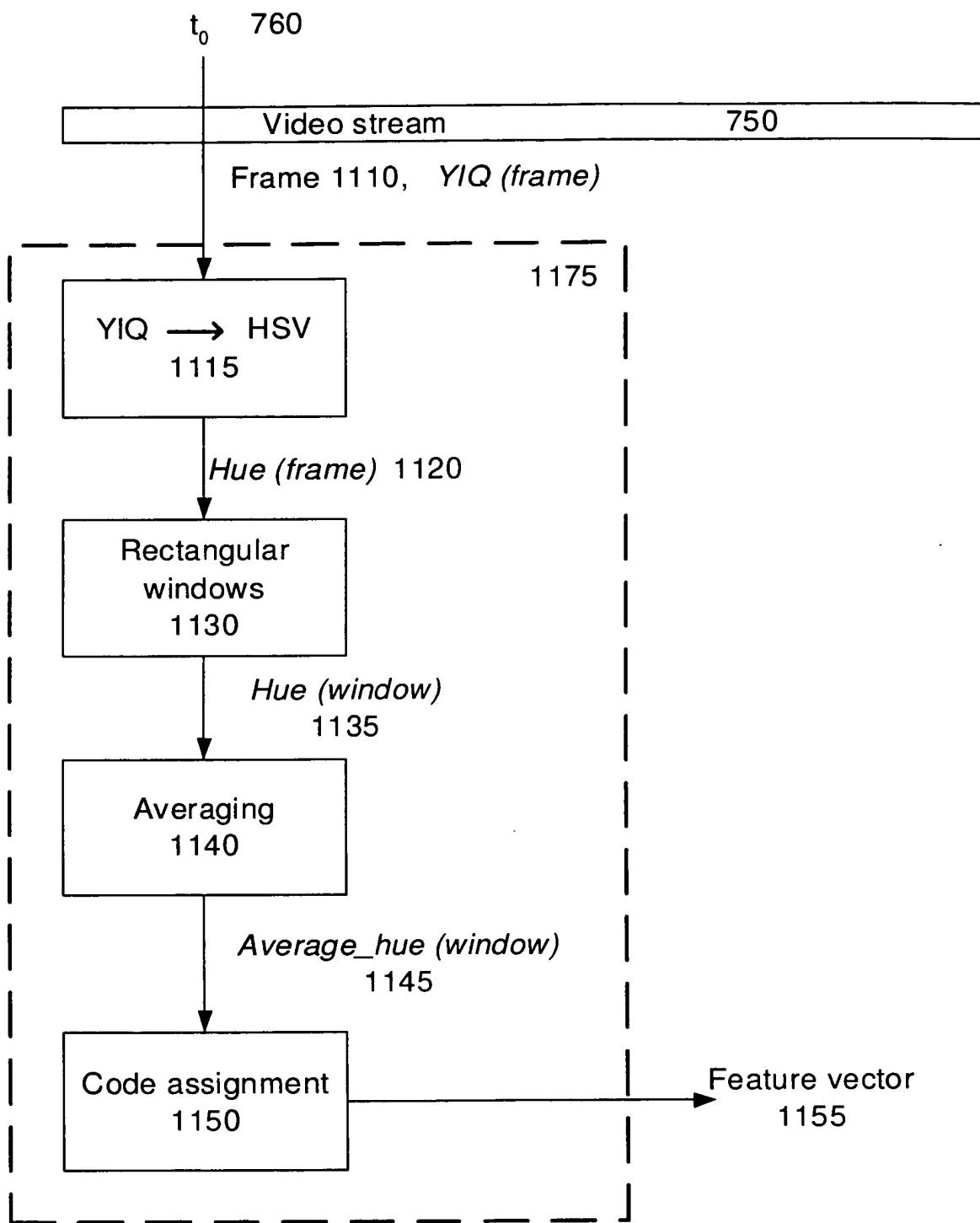
**Figure 9**



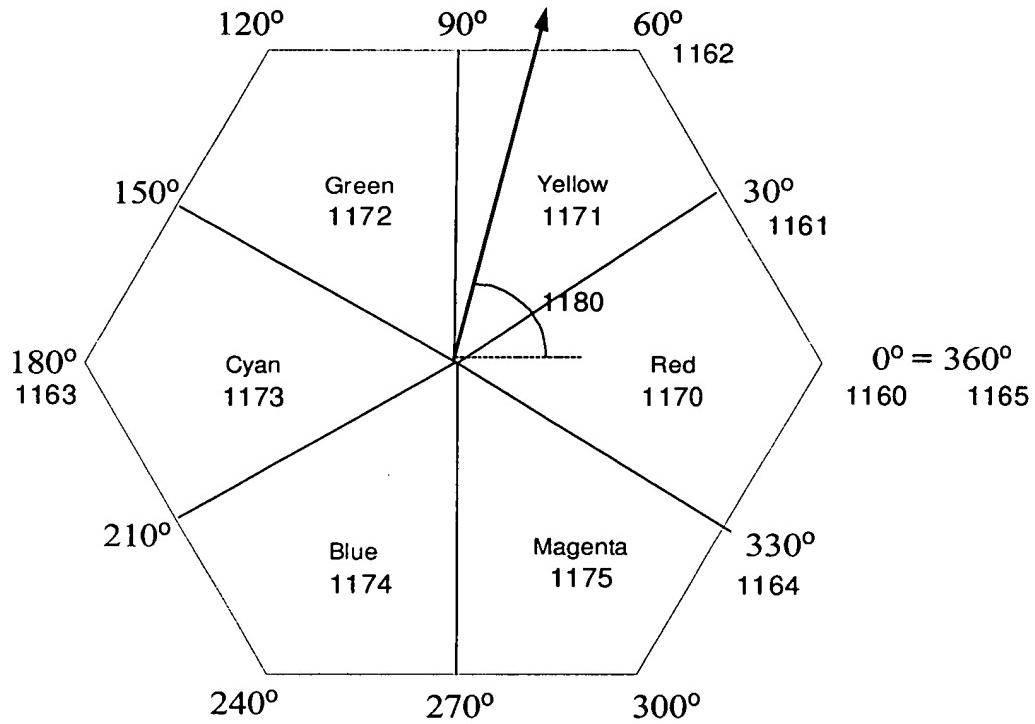
**Figure 9A**



**Figure 10**



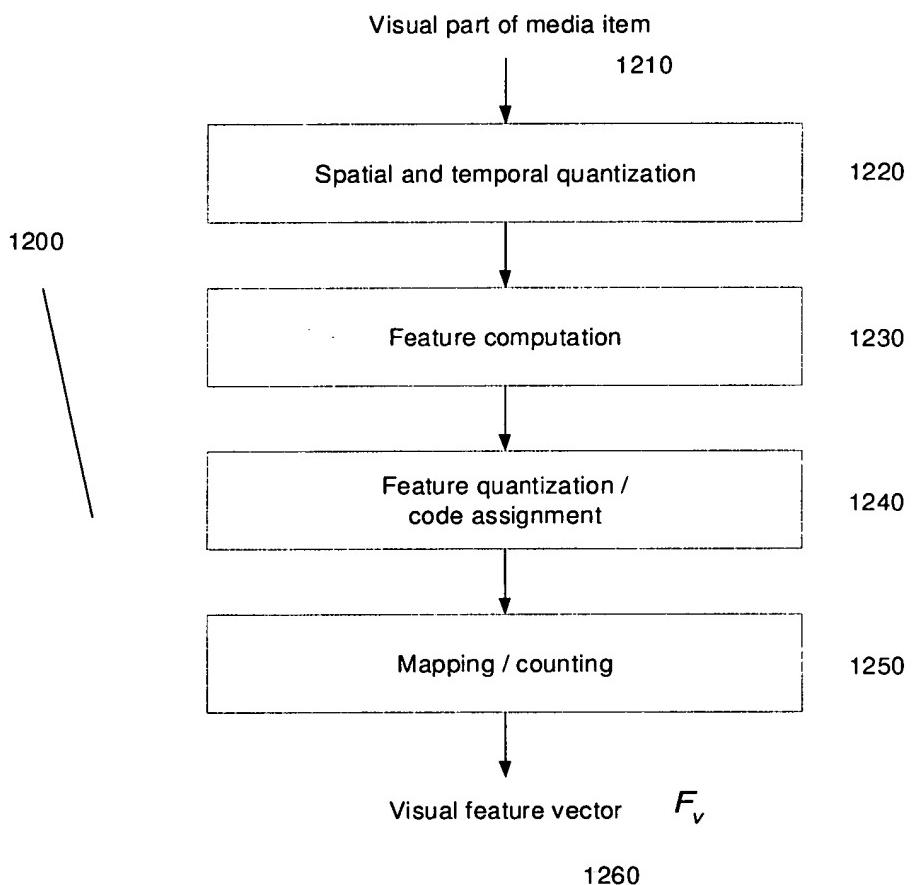
**Figure 11A**



Hue quantization (1180)

<i>Color range</i>		<i>Code</i>	<i>Color</i>	
330 - 30	degrees	0	Red	(1170)
30 - 90	degrees	1	Yellow	(1171)
90 - 150	degrees	2	Green	(1172)
150 - 210	degrees	3	Cyan	(1173)
210 - 270	degrees	4	Blue	(1174)
270 - 330	degrees	5	Magenta	(1175)

Figure 11B



**Figure 12**

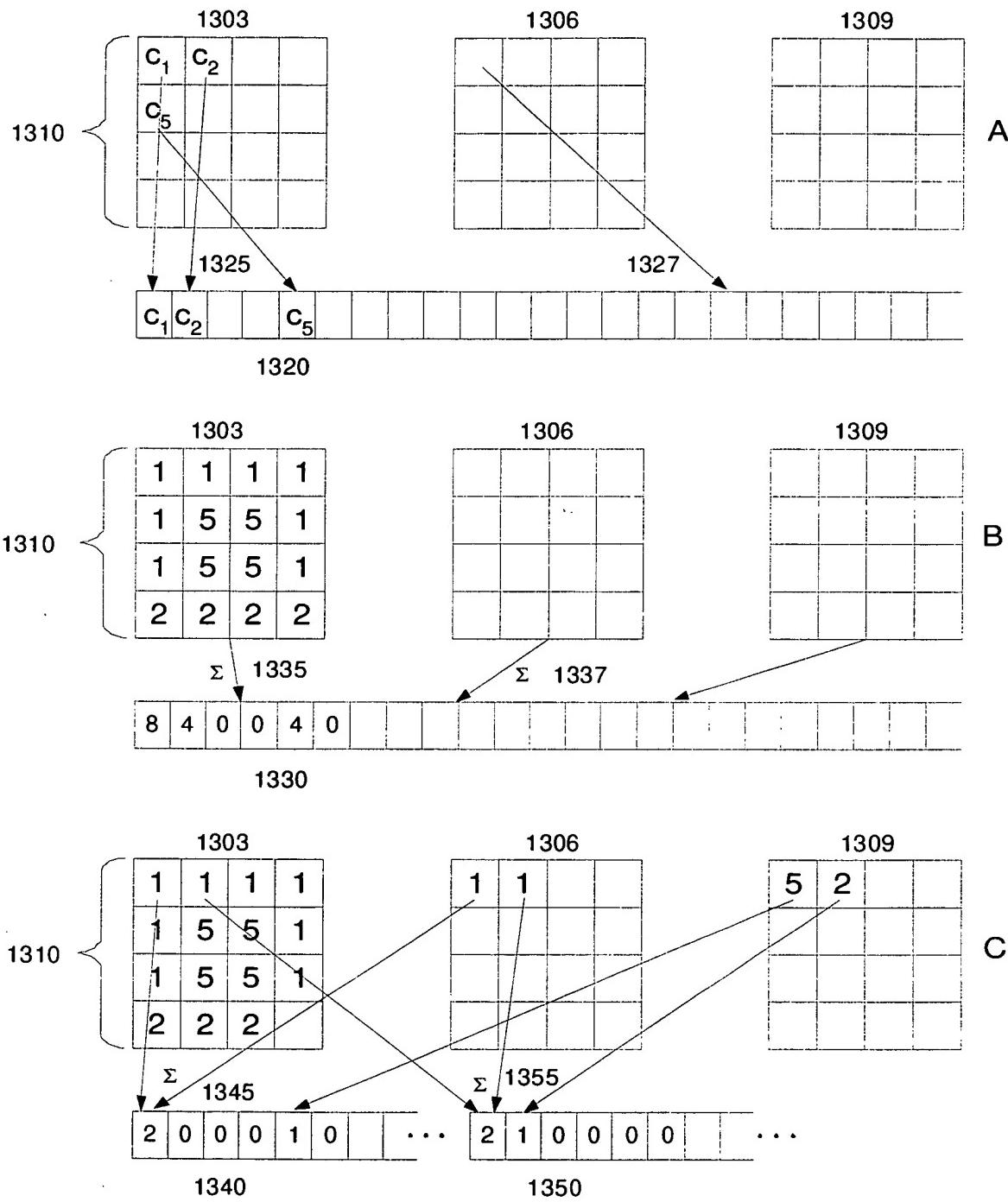
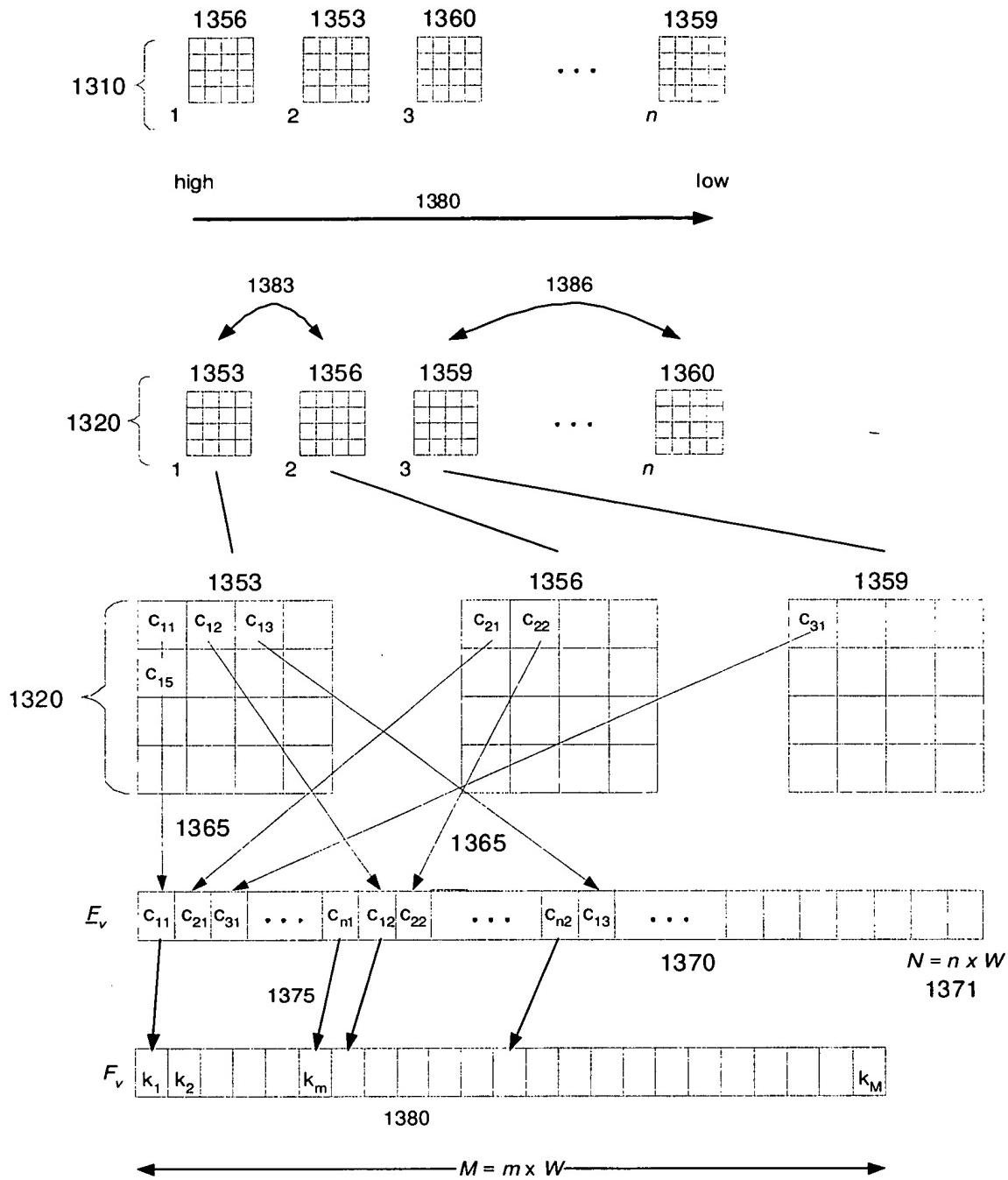


Figure 13



**Figure 13D**

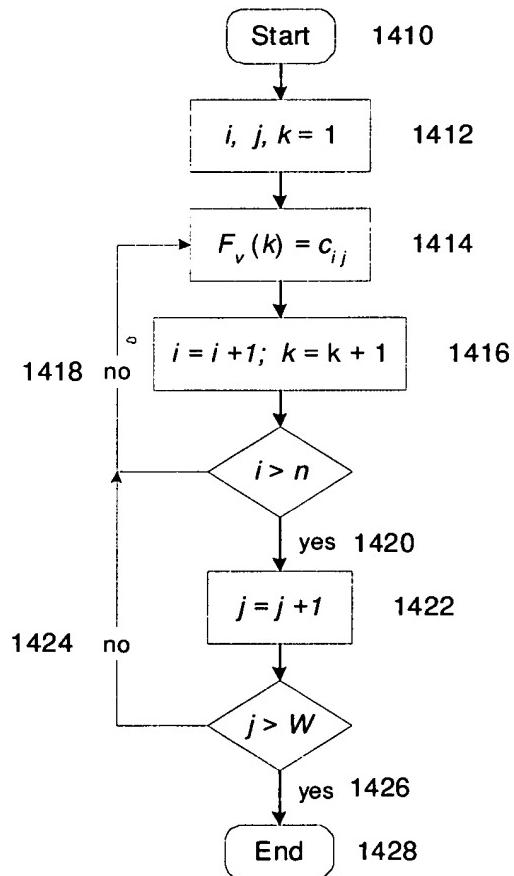
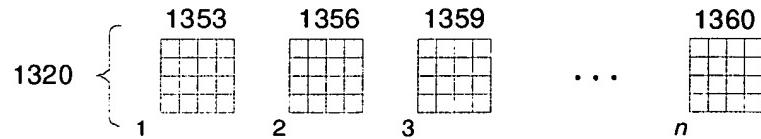
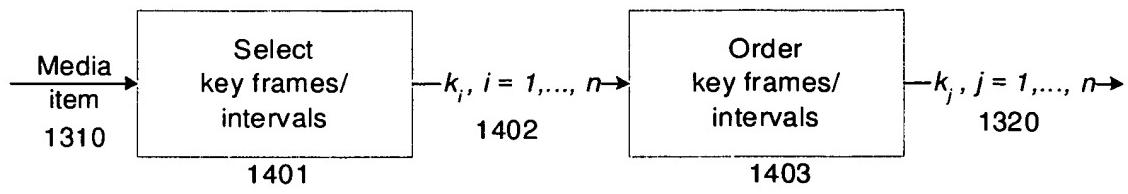


Figure 14A

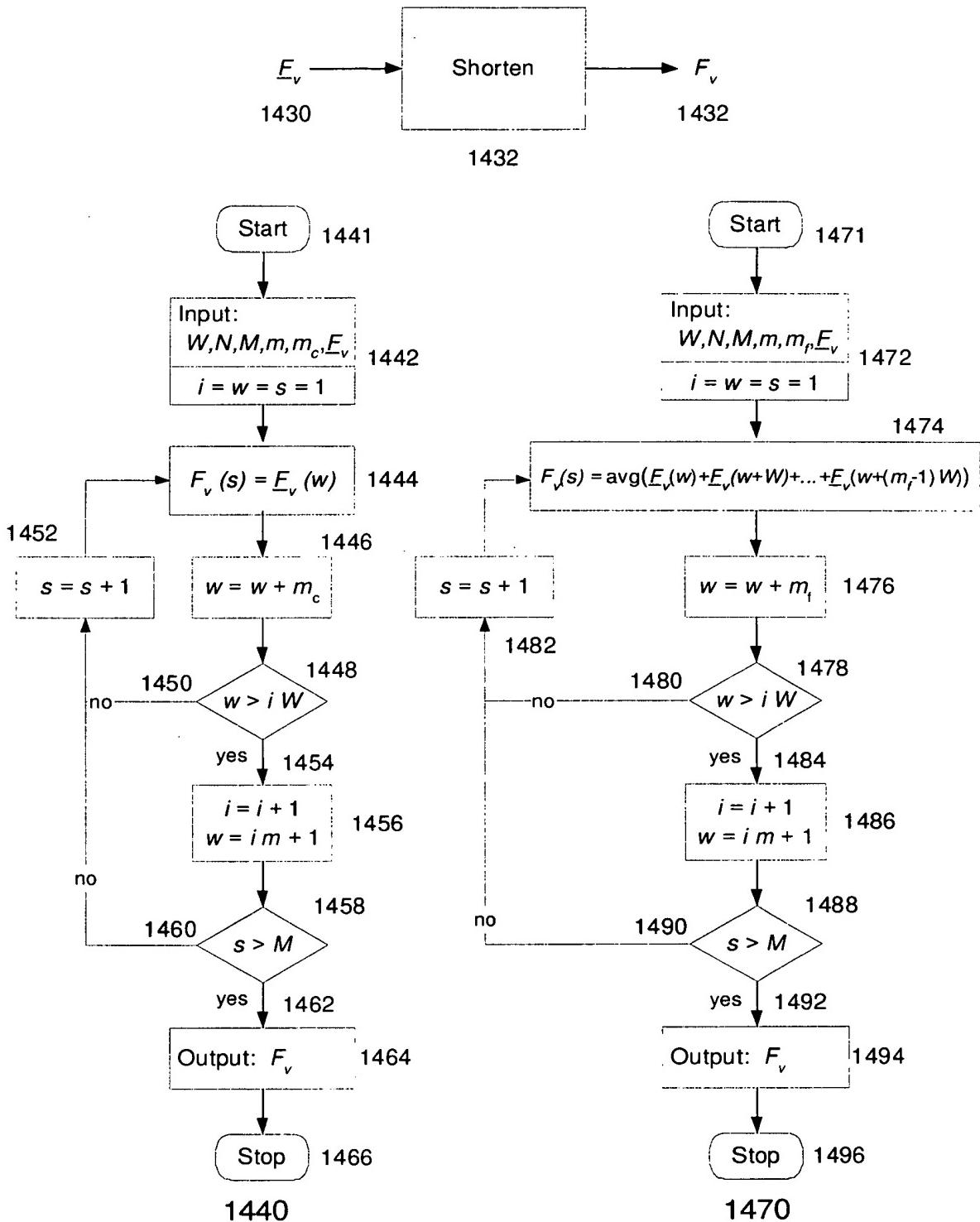
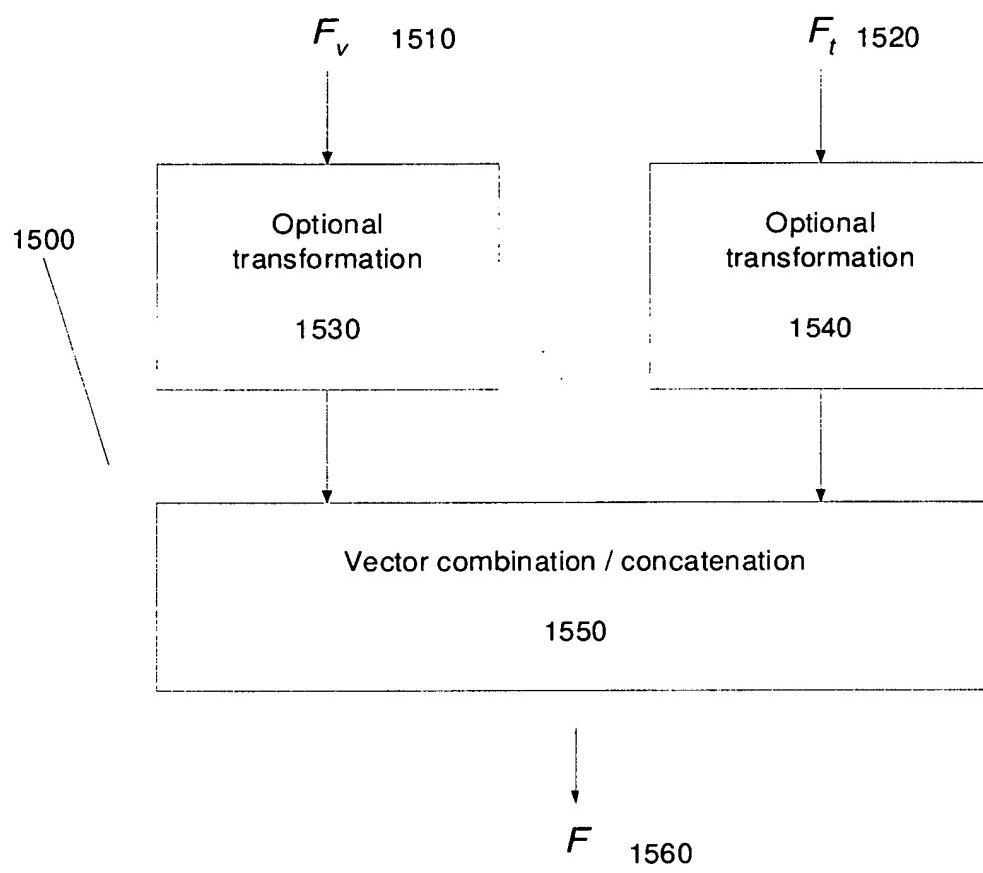
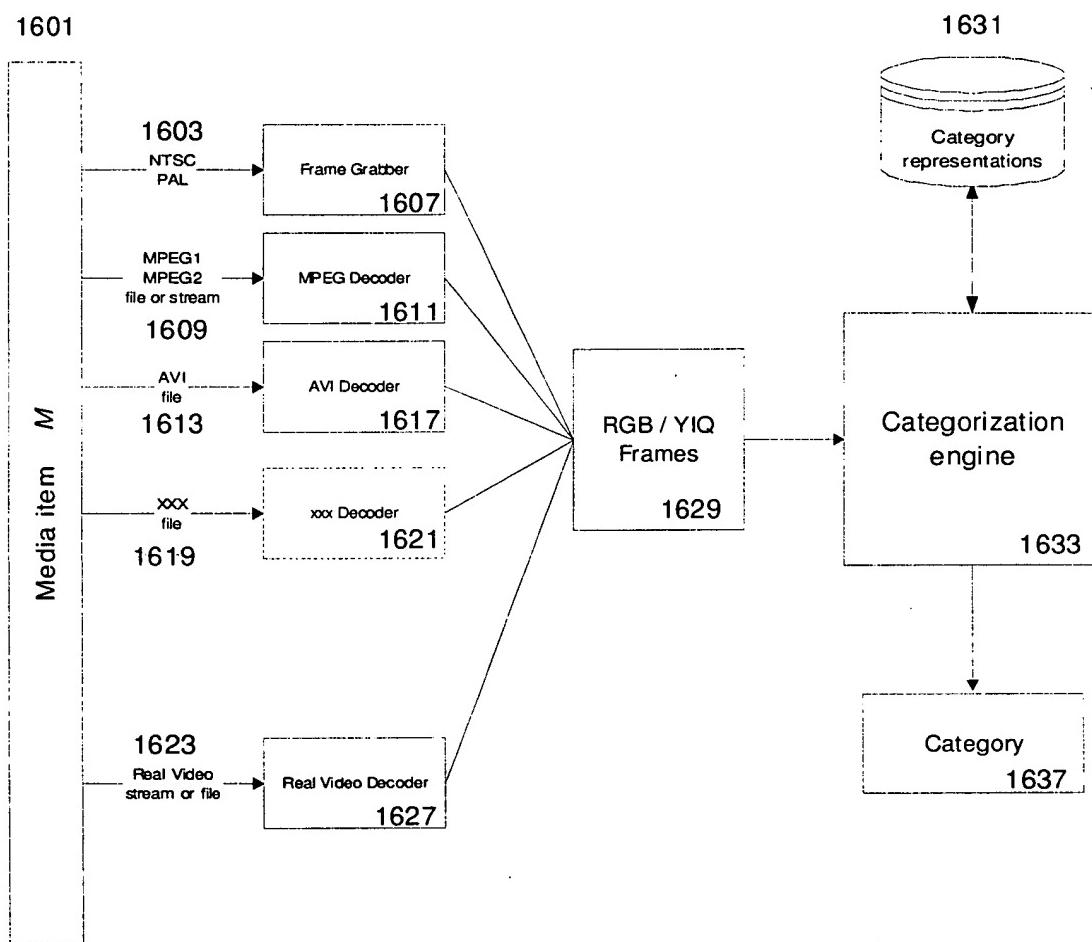


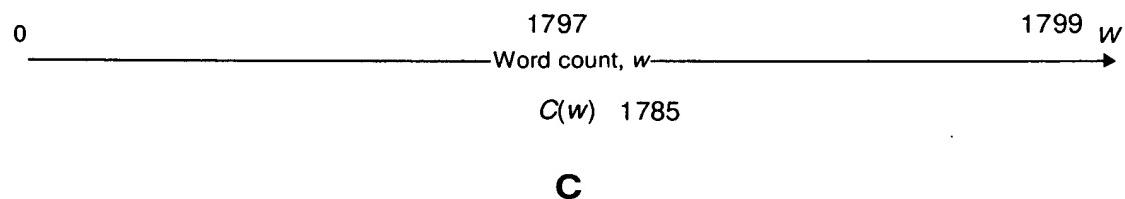
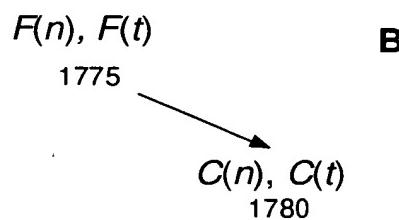
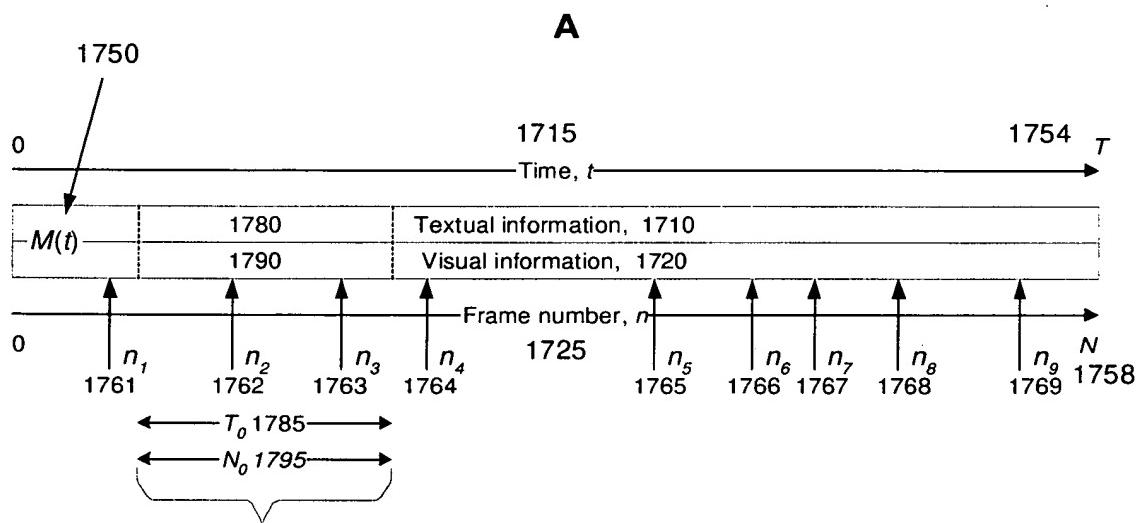
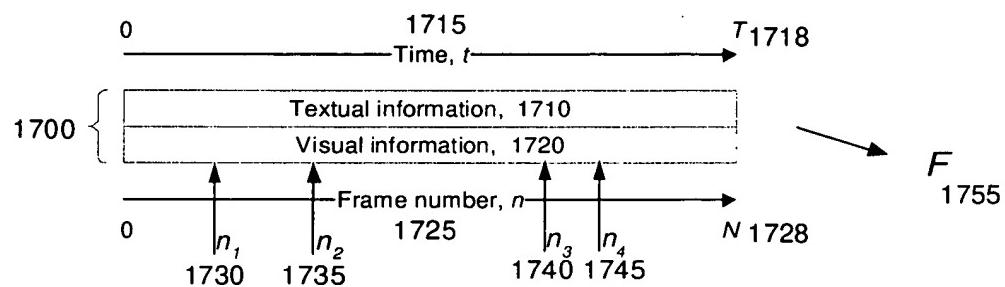
Figure 14B



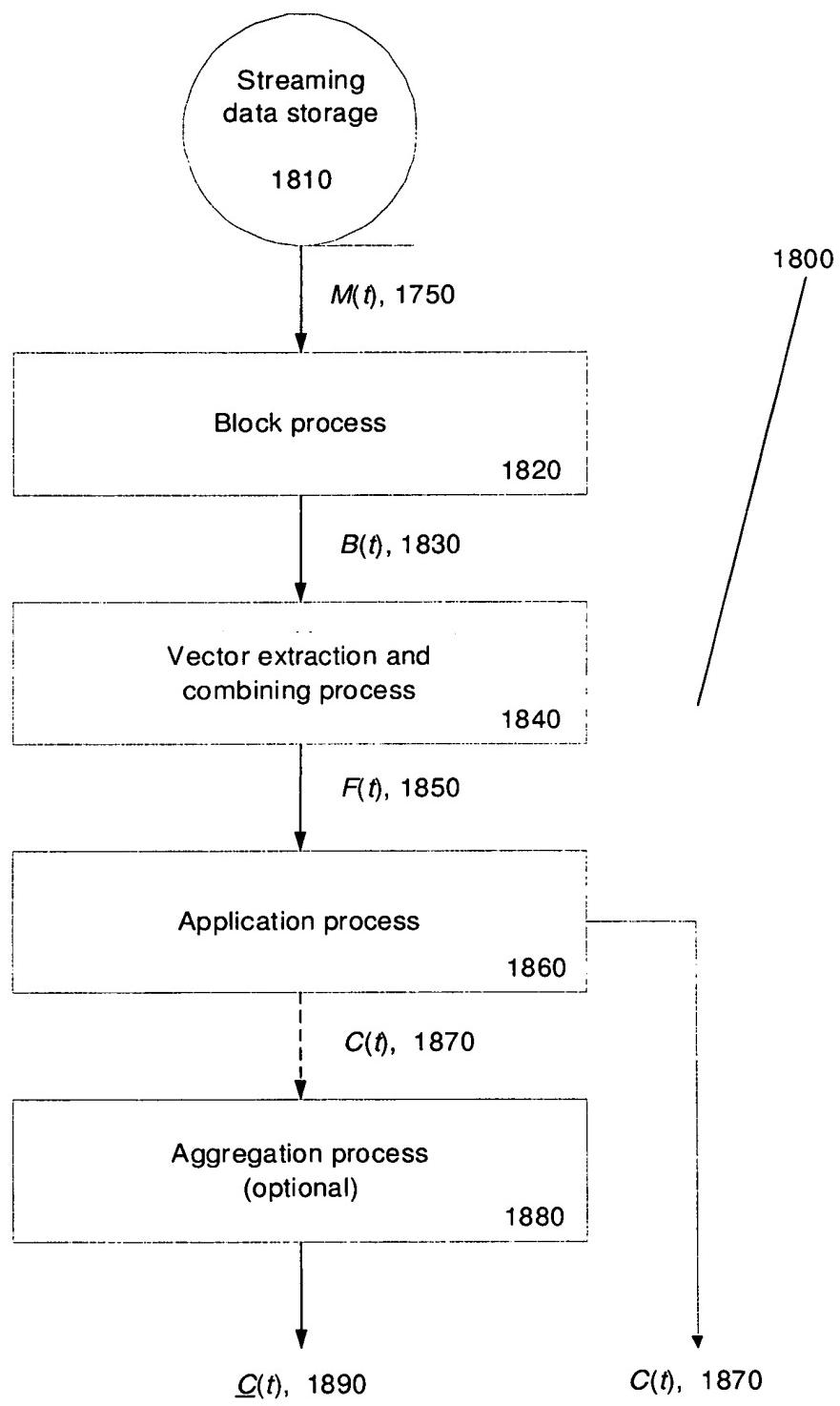
**Figure 15**



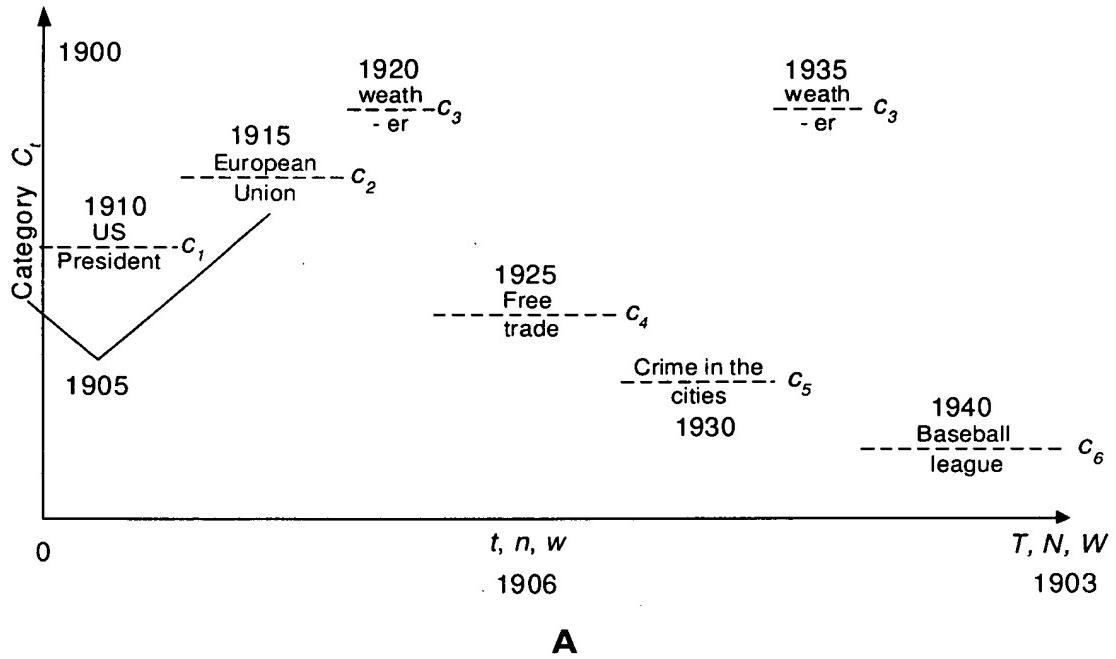
**Figure 16**



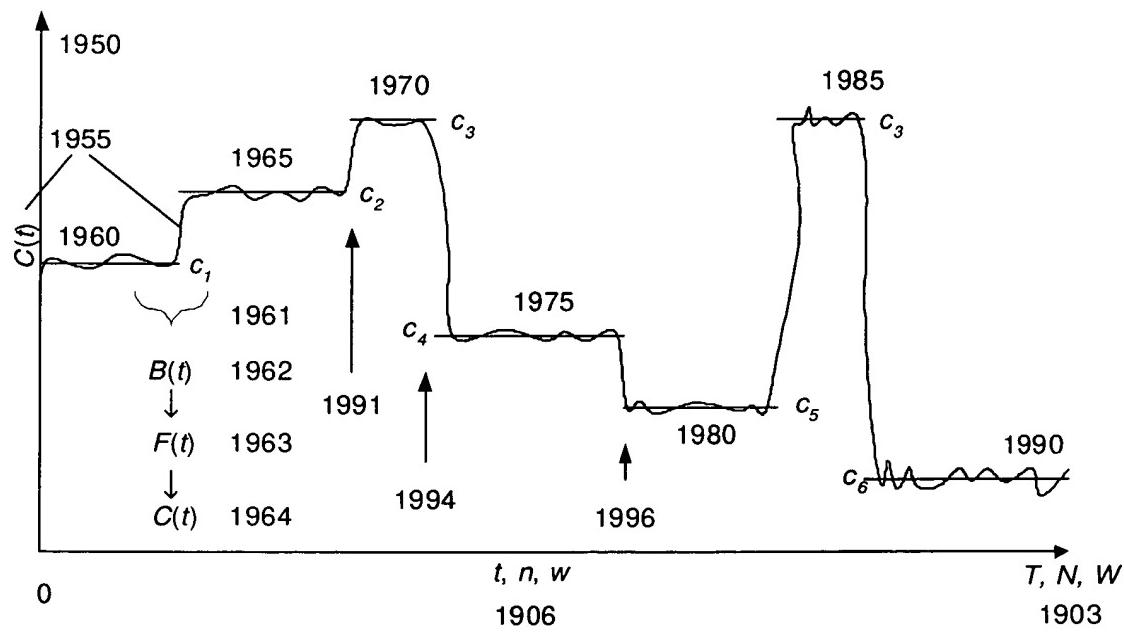
**Figure 17**



**Figure 18**

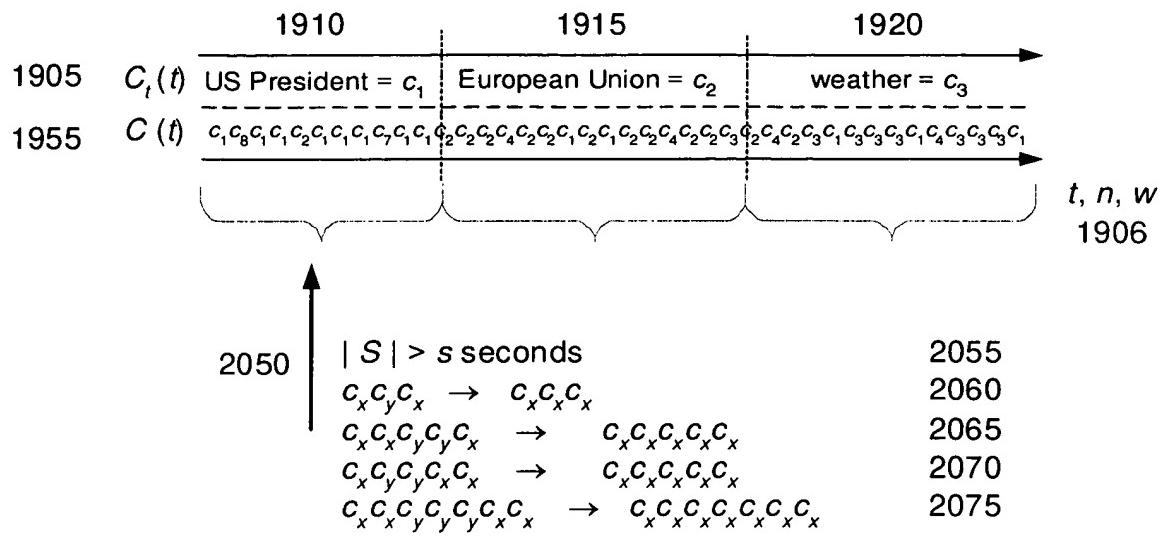


**A**

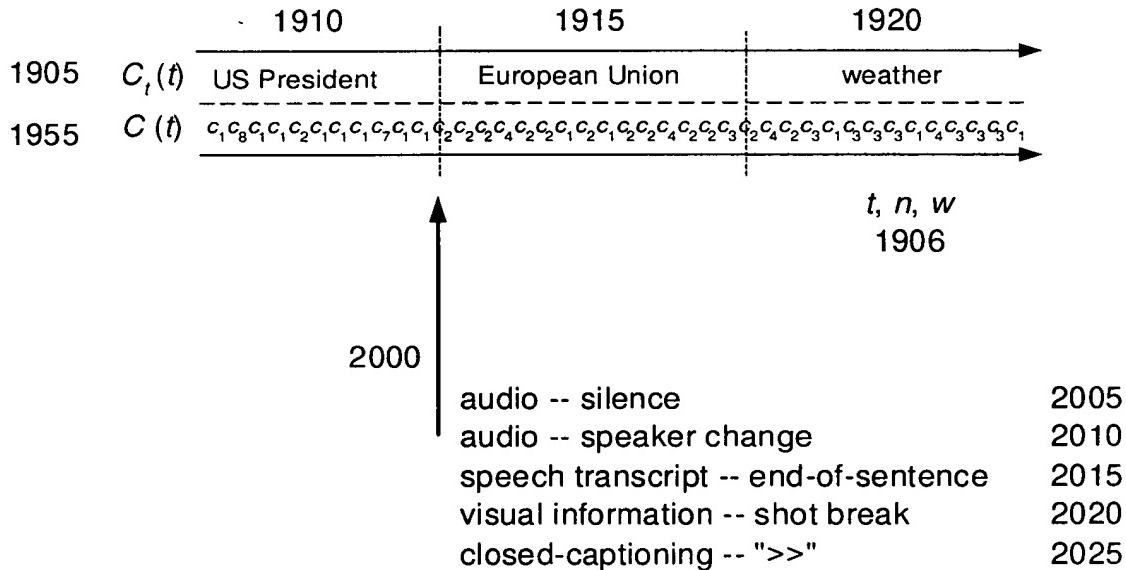


**B**

**Figure 19**



A



B

Figure 20

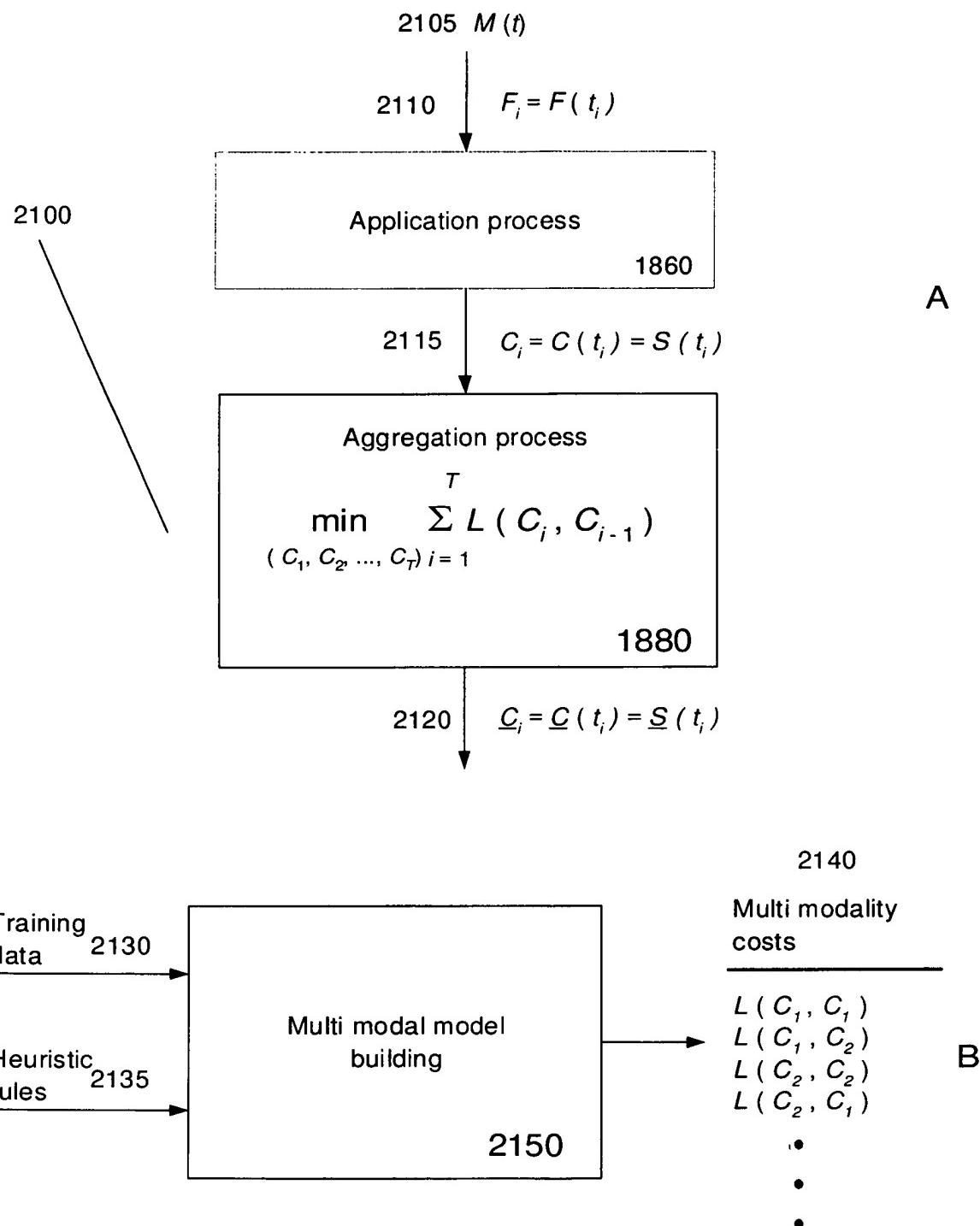


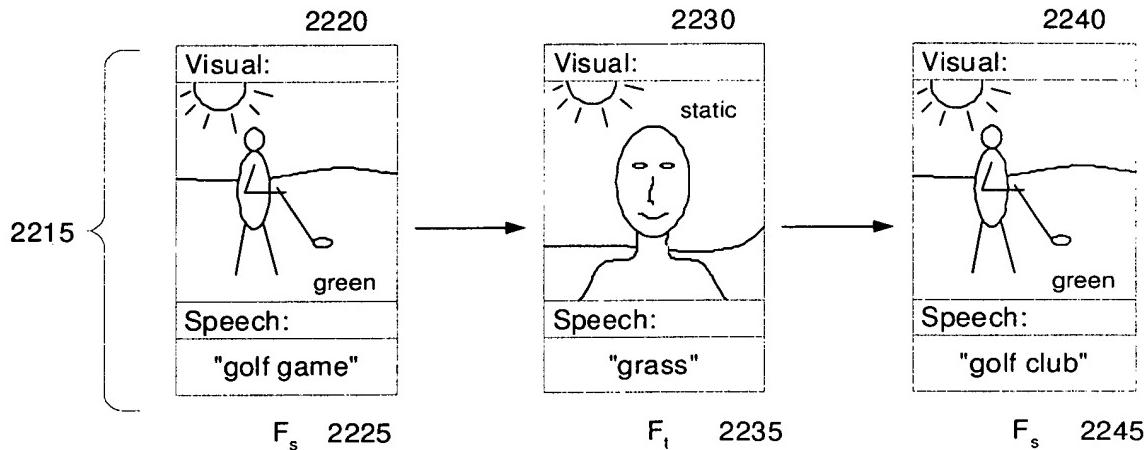
Figure 21

$F_s \rightarrow$  "sport"  
 $F_t \rightarrow$  "sport or disaster"  
 $F_d \rightarrow$  "disaster"

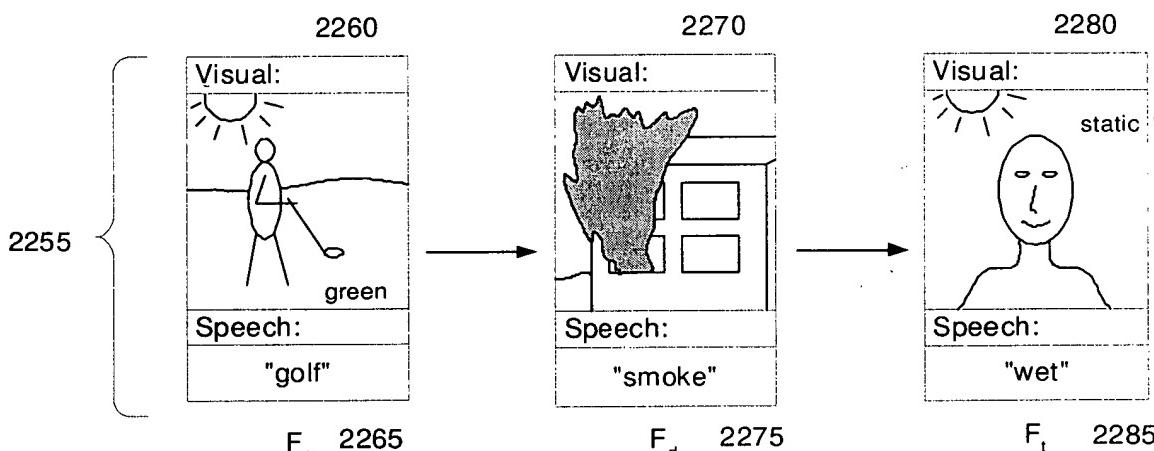
2205

$L(sport, sport)$	=	small
$L(sport, disaster)$	=	large
$L(disaster, sport)$	=	large
$L(disaster, disaster)$	=	small

2210



Interpretation:  
 sport  $\rightarrow$  sport  $\rightarrow$  sport  $\rightarrow$  sport 2250



Interpretation:  
 sport  $\rightarrow$  disaster  $\rightarrow$  disaster 2290

Figure 22